Correlating instrumentation data to system states: A building block for automated diagnosis and control by Cohen et al.

CSC 724 Paper review - Vaibhav Singh, vsingh7

January 29, 2019

1   Summary

The paper talks about the use of Tree-Augmented Bayesian Networks (TANs) to identify system level metrics and thresholds which can define system states as high or low performant.

2   Description

The paper presents an approach to creating an analysis engine to process the metrics and indicator values, and ultimately induce a classifier which predicts whether the system will fulfill SLOs or will have SLO violations over a period of time. The classifier also indirectly identifies the metrics which are most related to SLO violations.

This exercise leads to the metrics being modeled as a Bayesian network, specifically using TANs to reduce overfitting. The model is not chosen to be naive Bayes to ensure that events are not modeled to be mutually independent.

Overall balanced accuracy of TAN based models is high, from 87 to 94 per cent. A single metric alone (for ex, in a model created using naive Bayes) is insufficient to predict patterns of SLO violations. however, a small number of (mutually dependent) metrics are enough to predict SLO violations.

3   Strong Points

Models using TANs are easier to create than models using apriori knowledge. TAN models presented in the paper present strong correlation between a small number of metrics (3-8), leading to high accuracy of 90-95 per cent.
TAN models are easy to replicate and model, and are flexible enough to simulate real life scenarios well.

4   Weak Points

While the paper can be used for diagnosis and control, the results are too generic to be used to root cause issues.

The paper assumes most issues in a node to be caused by the node itself, and does not take into account the impact of different nodes in the system.

5 Improvement

The paper's ideas of creating TANs based on node level metrics should be enhanced by taking into account the impact of other nodes in the system as well.